Asynchronously Parallel Optimization Solver for Finding Multiple Minima

Jeffrey Larson & Stefan M. Wild

Mathematics and Computer Science Division

Argonne National Laboratory

U.S. Department of Energy

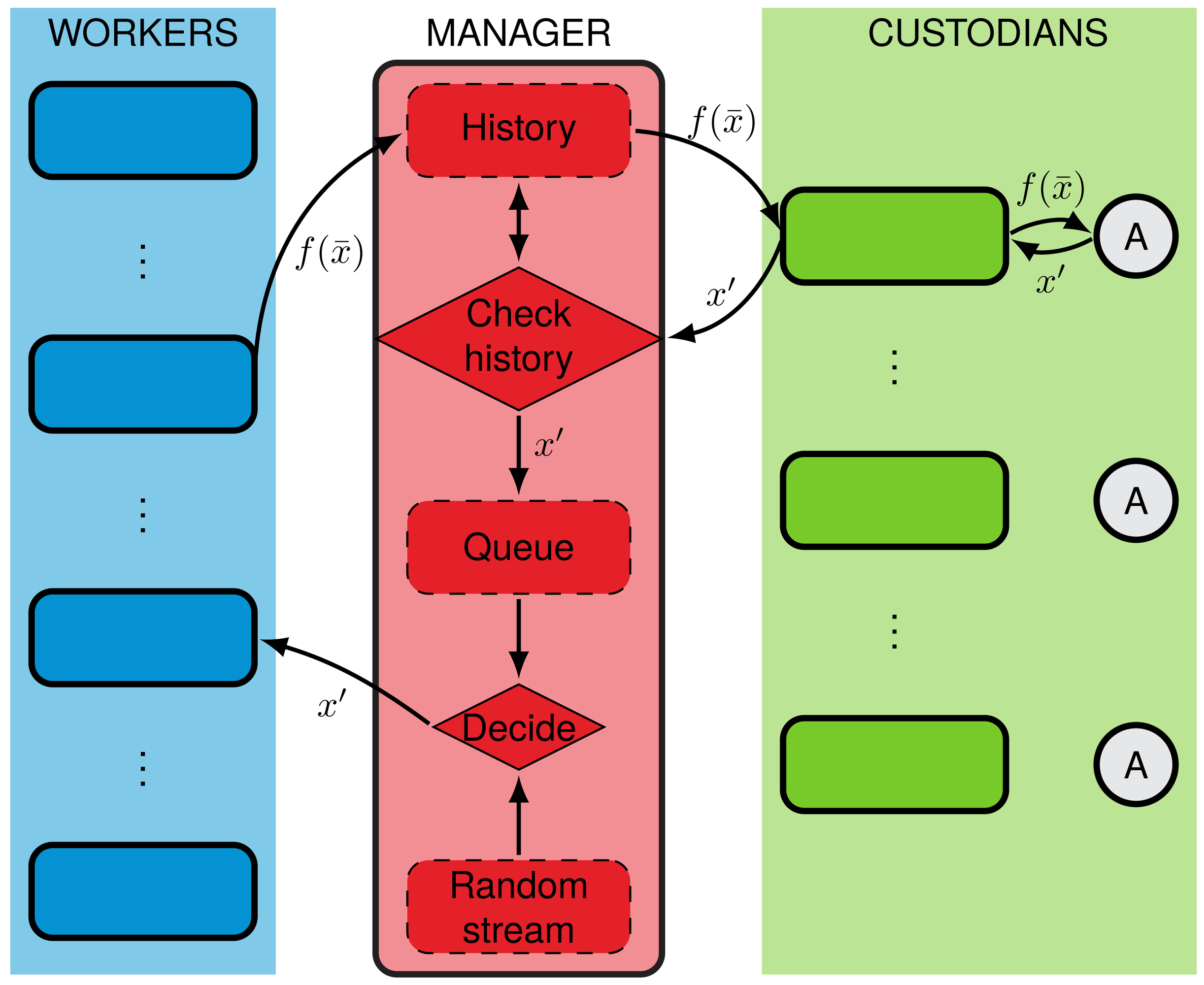
Problem statement

We want to find multiple, high-quality local minima of the nonlinear optimization problem

$$\underset{x\in\mathbb{R}^n}{\text{minimize}}\ f(x)$$
$$\text{subject to: } x \in \mathcal{D}$$

when $\mathcal{D}$ is compact, concurrent evaluations of f are possible, and relatively little is known about f.

Asynchronous workflow



Highlights

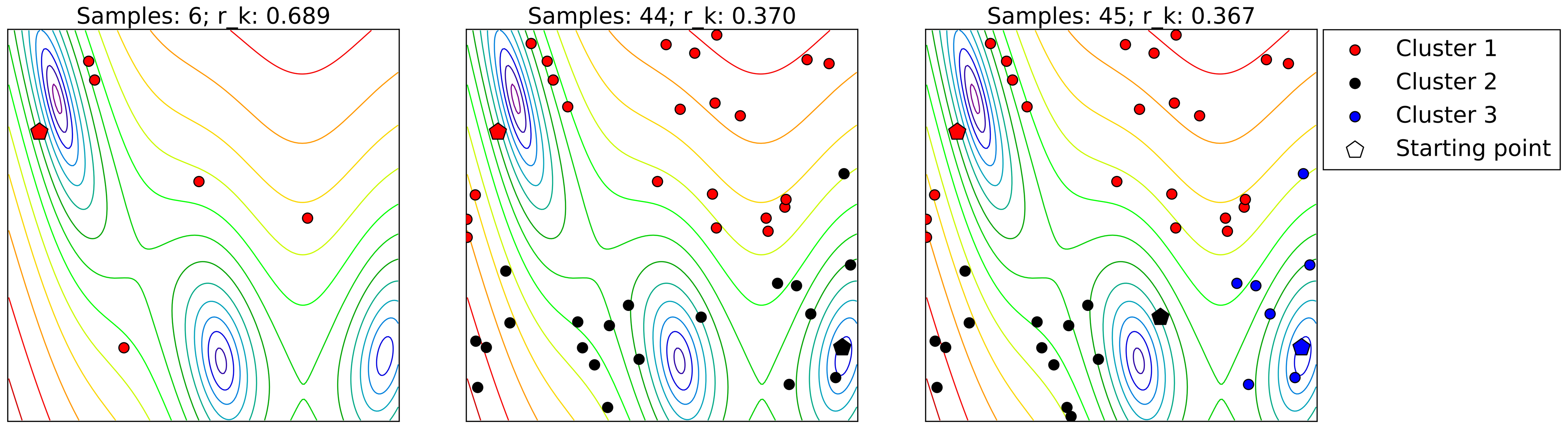
- Multistart algorithm that considers all previously evaluated points when deciding where to start or continue a local optimization run.
- Applicable to general optimization; its judicious use of function evaluations is especially suited for expensive derivative-free objectives.
- Time to solution scales well even when the time to evaluate the objective is highly variable.
- The algorithm has strong theoretical properties and performs well in practice.
- Depends on the critical distance

$$r_k = \frac{1}{\sqrt{\pi}} \sqrt[n]{\Gamma\left(1+\frac{n}{2}\right) \operatorname{vol}\left(\mathcal{D}\right) \frac{5\log|\mathcal{S}_k|}{|\mathcal{S}_k|}}. \quad (1)$$

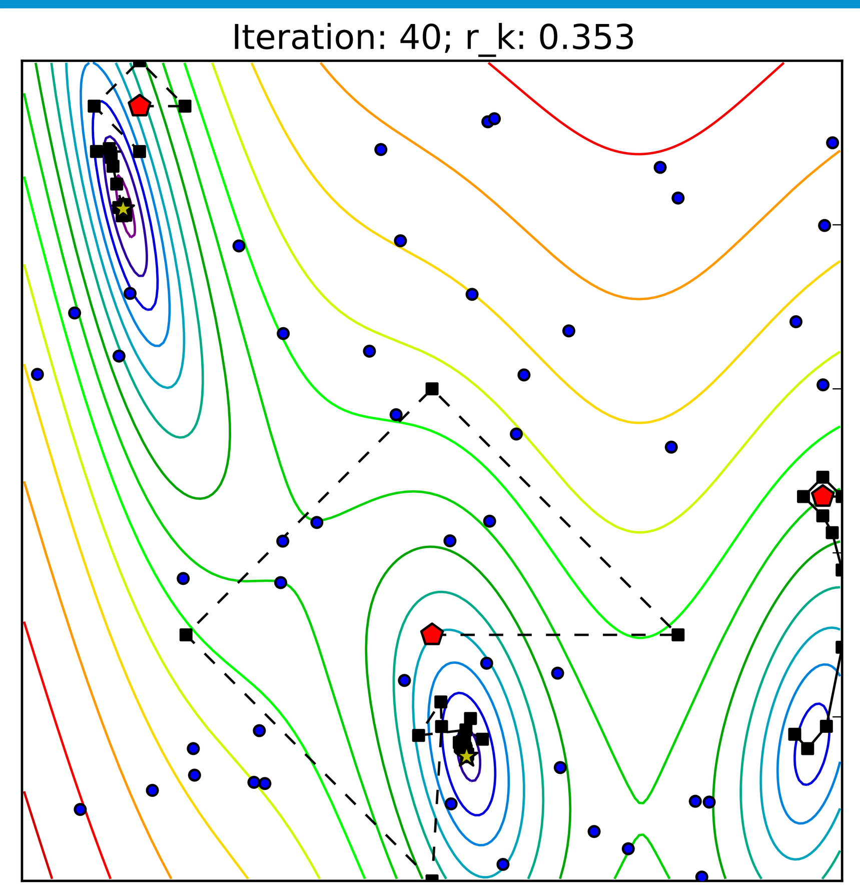
APOSMM

input: Local optimization method, random stream $\mathcal{R}_S$, tolerance ν.
for $w = \{1,\ldots,c\}$ **do**
 Give w a point from $\mathcal{R}_S$ at which to evaluate f.
for $k = 0,1,\ldots$ **do**
 Receive from worker w that has evaluated its point $\tilde{x}$.
 if $\tilde{x} \in A_k$ **then**
  if *$\tilde{x}$'s run is complete* **then**
   Add minimizer to X_k^*; remove points from run from A_k.
  else
   Query local optimization method and add subsequent point from $\tilde{x}$'s run to Q_L.
 Update H_k; Update r_k using (1).
 Start local optimization method at points in H_k satisfying certain conditions; add the subsequent point(s) to Q_L.
 Kill runs with candidate minima within 2ν of each other, keeping the best run.
 Give w a point x' at which to evaluate f, either from Q_L or $\mathcal{R}_S$.

Conditions

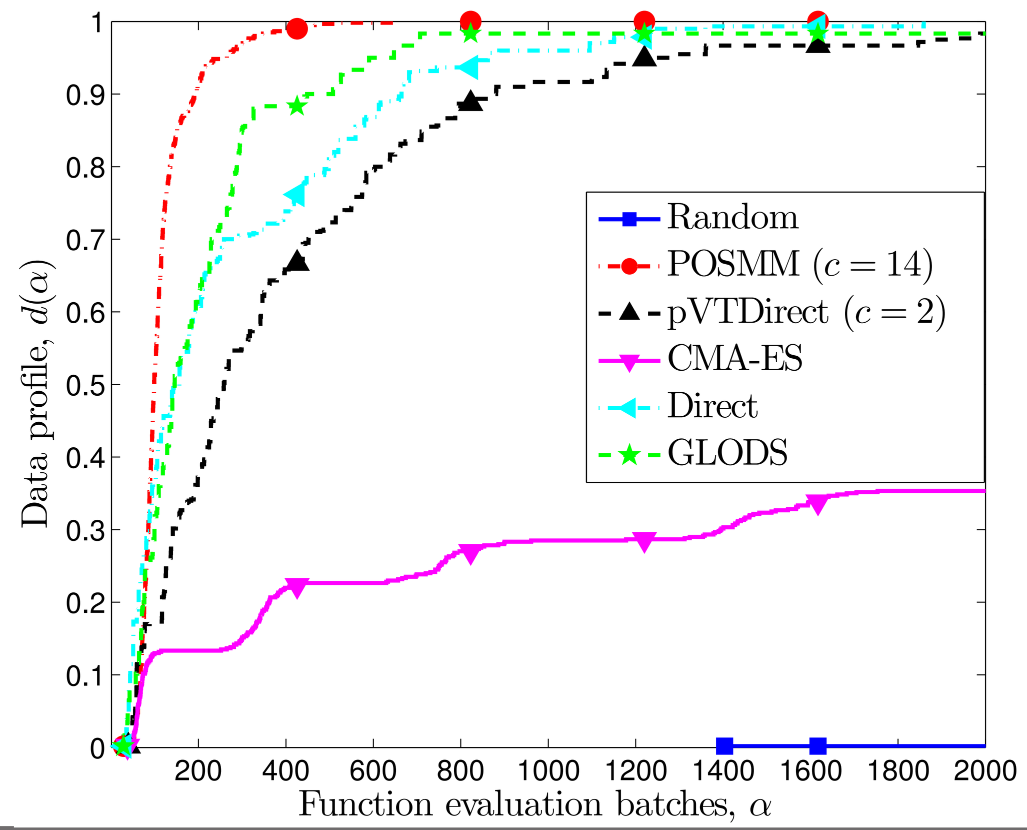


Snapshot



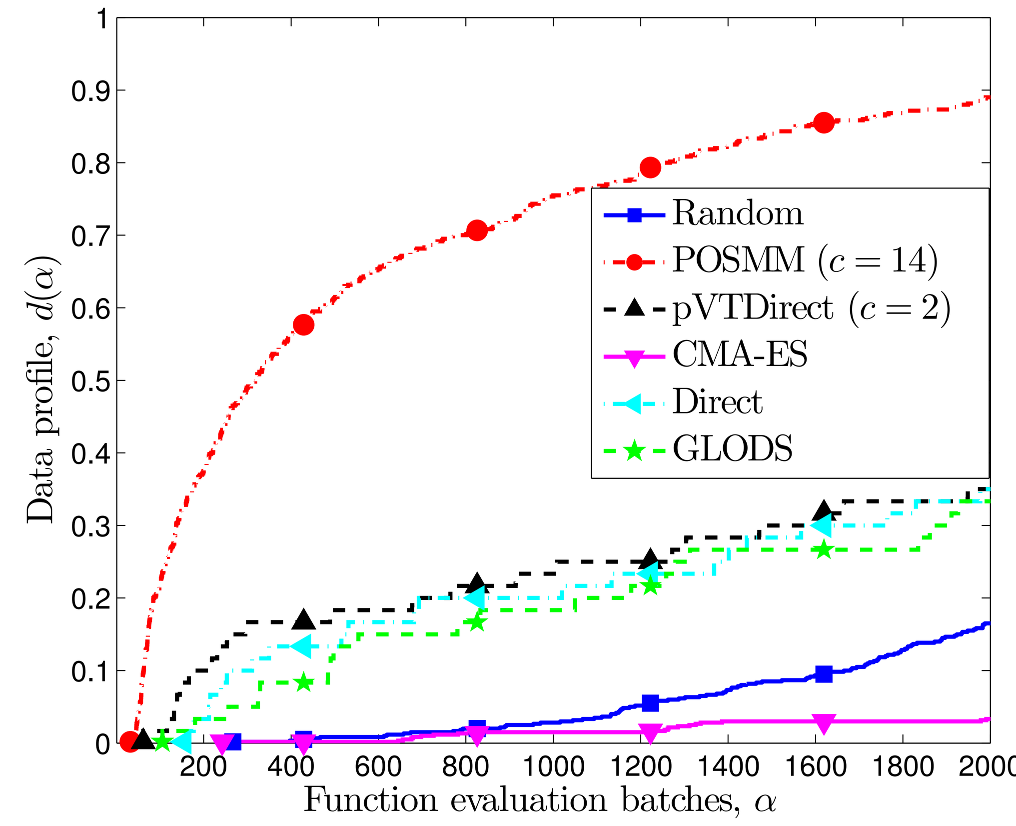
Approximating global minima

$$f(x) - f_G \le (1-10^{-5})\left(f(x_0) - f_G\right)$$

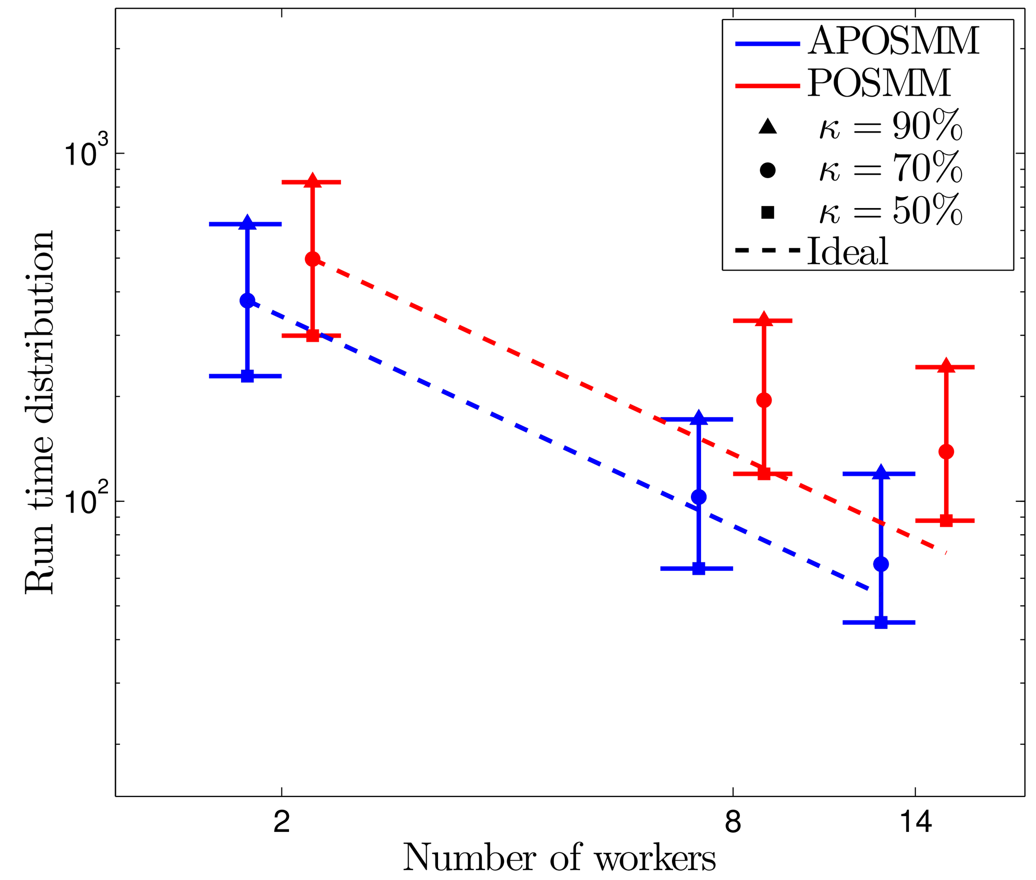


Finding multiple minima

Within $\sqrt[n]{\frac{10^{-4}\,\Gamma(\frac{n}{2}+1)}{\pi^{n/2}}}$ of 3 best minima



Performance scalability



Theorem

- If $f \in C^2$,
- there is a distance $\epsilon > 0$ between local minima,
- the local optimization method is strictly descent,
- r_k is defined by (1), then

APOSMM almost surely starts a finite number of local optimization runs and every local minimum is found or has a single local optimization run asymptotically converging to it.

J. Larson and S. M. Wild. Asynchronously parallel optimization solver for finding multiple minima. ANL/MCS-P5575-0316, 2016

J. Larson and S. M. Wild. A batch, derivative-free algorithm for finding multiple local minima. Optimization and Engineering 17(1), 2016